

REIHE INFORMATIK

20/95

**Mathematical Analysis of Computer Generated
Binary Fourier Transform Holograms**

E. Zhang, J. Hesser, C. Dietrich, S. Nochte, R. Männer

Universität Mannheim

Seminargebäude A5

D-68131 Mannheim

Mathematical Analysis of Computer Generated Binary Fourier Transform Holograms

Eryi Zhang, Jürgen Hesser, Christoph H. Dietrich, Steffen Noehte, Reinhard Männer

Chair for Computer Science V, University of Mannheim, D-68131 Mannheim, Germany

A rigorous mathematical treatment of a general binarization process for computer generated binary Fourier transform hologram is developed. Further a generalized error diffusion coefficient matrix is derived. In the view of these mathematical results we distinguish different binarization methods by the error diffusion process. Some binarization methods, their properties, and computational demands are investigated. Examples are presented.

1. Introduction

Superior characteristics of binary holograms in terms of ease of production [1,2] and lower noise sensitivity [3], compared to gray-scale holograms, have led to an increased research over the last years.

Coherent light diffracted by holograms forms a reconstruction wave. We are particularly interested in a certain sub-region in the reconstruction, which we call reconstruction region. Different methods exist to generate binary holograms, where the emphasis of exactness lies only on the reconstruction region; the surrounding area may have arbitrary reconstruction errors.

While hard clipping uses a constant threshold for binarization, error diffusion additionally diffuses the binarization error locally [5-10]. Direct binary search [11-13] starts with a randomly generated binary hologram. Iteratively, one pixel value at a time is flipped where the flip is accepted only if the error in the reconstruction region is reduced. Iterative binarization [14-16] is based on repeating the following steps: From a partly binarized hologram a reconstruction is calculated. In the reconstruction, the reconstruction region is replaced by the original object, and from the modified reconstruction a new hologram is computed. Each step leads to a reduction of errors in the reconstruction region.

In this paper a general binarization process for binary Fourier transform hologram is developed. A generalized error diffusion coefficient matrix is derived. Section 2 describes the mathematical framework that details the binarization process. In section 3, properties of the generalized error diffusion coefficients are investigated. Binarization noise appeared in the reconstruction region is estimated in section 4. These results are applied to noniterative algorithm, error diffusion, and to iterative algorithm, which are described in section 5 and section 6. Some generated binary holograms and their simulated reconstruction are presented.

2. Mathematical foundations of the binarization process

Starting from a desired object wave $u(x,y)$, its diffraction wave $U(\mu,v)$ on the hologram plane is calculated by a digital Fourier transformation. We use capital letters to denote quantities in Fourier transform domain, corresponding small letters denote quantities in object (or reconstruction) domain. $u(x,y)$ is commonly modified with certain pseudo-phase model to reduce the dynamic range of $U(\mu,v)$. As in experimental holography, a plane reference wave $R(\mu,v)=\exp\{+j2\pi(x_0\mu+y_0v)\}$ is introduced, where (x_0,y_0) determines the direction of the reference wave. $R(\mu,v)$ is then superposed on the diffraction wave $U(\mu,v)$ to form a positive real-valued interference pattern $H(\mu,v)$:

$$H(\mu,v)=\mathcal{R}(|U(\mu,v)|\cos[\Phi(\mu,v)-2\pi(x_0\mu+y_0v)]-Bs) \quad (1)$$

where $U(\mu, \nu) = |U(\mu, \nu)| \exp\{j\Phi(\mu, \nu)\}$ is normalized to 1, B_s is a constant bias to ensure $H(\mu, \nu) \geq 0$, $B_s = \min\{|U(\mu, \nu)| \cos[\Phi(\mu, \nu) - 2\pi(x_0\mu + y_0\nu)]\}$, \mathcal{N} is normalization factor. Thus, it is ensured that $0 \leq H(\mu, \nu) \leq 1$.

If the interference pattern $H(\mu, \nu)$ is exposed onto a photo material, e.g. a film or plate as amplitude transmission, then an amplitude Fourier transform hologram is obtained.

A common Fourier transform configuration is used for the replay of the hologram. That is, the hologram $H(\mu, \nu)$ is located in the front focal plane of the system, and illuminated by a collimated wave parallel to the optical axis. The reconstruction $h(x, y)$ appears in the back focal plane, which contains the object wave $u'(x, y)$ off axis with an offset (x_0, y_0) , its conjugate, and a delta function in the center. The offset (x_0, y_0) must be large enough to separate $u'(x, y)$ from both its twin image and the delta function. If $H(\mu, \nu)$ is analog valued and the amplitude transmission of the photo material is proportional to $H(\mu, \nu)$, then the reconstruction $u'(x, y)$ is identical to the original object wave $u(x, y)$, i.e., $u'(x+x_0, y+y_0) = u(x, y)$. Quantization of $H(\mu, \nu)$ leads to an error of $u'(x, y)$ in respect to $u(x, y)$. This error mainly depends on the number of quantization levels. Obviously the quantization error is the largest for binary holograms. To minimize this error, many methods have been proposed as mentioned in introduction.

In the following a general relation between the binarized pixels in the hologram and the resulting error in the reconstruction region is inferred. Without loss of generality, we assume that the hologram $H(\mu, \nu)$ has $N \times N$ pixels, and the reconstruction region w in $h(x, y)$ has $M_x \times M_y$ ($0 < M_x, M_y < N/2$) pixels. That is, in hologram domain we have a 2-dimensional amplitude gray-scale hologram $H(\mu, \nu)$ given by Eq. (1) with $N \times N = N^2$ pixels ($\mu, \nu = 0, 1, 2, \dots, N-1$); and in object domain we want to have the function $u(x, y)$ reconstructed by a binarized $H(\mu, \nu)$ in the reconstruction region w . This region has $M_x \times M_y$ pixels and is centered at (x_0, y_0) . We use the Mean-Squared Error (MSE) to measure the difference between the desired object $u(x, y)$ and the reconstructed one. The latter is contained in the reconstruction $h(x, y)$ in region w :

$$MSE = \frac{1}{M_x M_y} \sum_{(x, y) \in w} |u(x - x_0, y - y_0) - \alpha h(x, y)|^2 \quad (2)$$

where α is a scaling factor.

MSE reaches the minimum if $\alpha = \frac{\sum_w u(x - x_0, y - y_0) h^*(x, y)}{\sum_w |h(x, y)|^2}$, where $h^*(x, y)$ is the conjugate of $h(x, y)$.

Substituting α into (2) and normalizing (2) by $\sum_w |u(x - x_0, y - y_0)|^2$, we arrive at

$$MSE = 1.0 - \frac{|\sum_w u(x - x_0, y - y_0) h^*(x, y)|^2}{\sum_w |u(x - x_0, y - y_0)|^2 \sum_w |h(x, y)|^2} \quad (3)$$

Thus, MSE has value range $0 \leq MSE \leq 1$. Note that MSE calculated by (2) or (3) is independent of the actual values of $h(x, y)$. That is, $MSE = 0$ if $h(x, y)$ in region w is only an amplification of the magnitude and/or a constant phase shift of the original signal. In the following we use the normalized MSE, Eq. (3).

Our aim is to generate a binary Fourier transform hologram, with that to reproduce the desired signal $u(x, y)$ in the reconstruction region as well as possible. In other words, the aim is to minimize MSE.

We perform the binarization process as follows:

First, we binarize K ($1 \leq K < N^2$) pixels in the gray-scale hologram. These K pixels are chosen uniformly and randomly. Then we observe their effect on the reconstruction by simulating the reconstruction of the hologram, and evaluating MSE in the reconstruction region with respect to the desired object wave. Non-zero MSE is resulted due to the binarization of the K pixels in the hologram. Subsequently, we correct pixel values in the reconstruction region by eliminating binarization error in this region. Finally, from this modified reconstruction a new

hologram is computed. Thus, the new hologram is connected with the partly binarized hologram by the correction of the reconstruction region.

This binarization process is represented mathematically as follows:

Let $H(\mu, \nu) = H^{(1)}(\mu, \nu)$ be the original gray-scale hologram. The upper index indicates the step number of the iterative process. And let the binarization error in the hologram be $E^{(1)}(\mu, \nu)$ for $(\mu, \nu) = (\mu_1, \nu_1), (\mu_2, \nu_2), \dots, (\mu_K, \nu_K)$:

$$E^{(1)}(\mu, \nu) = \begin{cases} S^{(1)}(\mu, \nu) - H^{(1)}(\mu, \nu) & \text{if } (\mu, \nu) = (\mu_1, \nu_1), (\mu_2, \nu_2), \dots, (\mu_K, \nu_K); \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$\text{where } S^{(1)}(\mu, \nu) = \begin{cases} 0 & \text{if } H^{(1)}(\mu, \nu) < T \\ 1 & \text{if } H^{(1)}(\mu, \nu) \geq T \end{cases}$$

$$(\mu, \nu) = (\mu_1, \nu_1), (\mu_2, \nu_2), \dots, (\mu_K, \nu_K).$$

Here T is a constant threshold. $S^{(1)}(\mu, \nu)$ is the binarized pixels with value of $\{0, 1\}$.

After the binarization we can write the new hologram transmission as:

$$H^{(1)}(\mu, \nu) = H^{(1)}(\mu, \nu) + E^{(1)}(\mu, \nu) \quad \mu, \nu = 0, 1, 2, \dots, N-1.$$

The operation above introduces errors in the reconstruction inside and outside the region w , i.e., the original signal with additional noise is obtained:

$$\begin{aligned} h^{(1)}(x, y) &= I\mathcal{F}\mathcal{T}\{H^{(1)}(\mu, \nu)\} \\ &= h^{(1)}(x, y) + e^{(1)}(x, y) \end{aligned} \quad x, y = 0, 1, 2, \dots, N-1,$$

where $I\mathcal{F}\mathcal{T}$ is the Inverse Fourier Transform, $h^{(1)}(x, y)$ is the reconstruction of the gray-scale hologram $H^{(1)}(\mu, \nu)$, $e^{(1)}(x, y) = I\mathcal{F}\mathcal{T}\{E^{(1)}(\mu, \nu)\}$ is error in the reconstruction caused by the binarization of $H^{(1)}(\mu, \nu)$. We call $e^{(1)}(x, y)$ binarization noise in the reconstruction. Note that $e^{(1)}(x, y)$ spreads over the whole reconstruction plane, $h^{(1)}(x, y)$ contains $u(x, y)$ in region w without error.

To improve the reconstruction in region w , we eliminate binarization noise only in this region. Thus, the signal left is identical to that reconstructed by the original gray-scale hologram. This operation is performed in region w , as well as in its conjugate region:

$$\begin{aligned} h^{(2)}(x, y) &= h^{(1)}(x, y) - \text{Rect}\left(\frac{x-x_0}{M_x}, \frac{y-y_0}{M_y}\right) e^{(1)}(x, y) - \text{Rect}\left(\frac{x+x_0}{M_x}, \frac{y+y_0}{M_y}\right) e^{(1)}(x, y) \\ &= h^{(1)}(x, y) + e^{(1)}(x, y) - \left[\text{Rect}\left(\frac{x-x_0}{M_x}, \frac{y-y_0}{M_y}\right) + \text{Rect}\left(\frac{x+x_0}{M_x}, \frac{y+y_0}{M_y}\right) \right] e^{(1)}(x, y) \end{aligned} \quad (5)$$

$$x, y = 0, 1, 2, \dots, N-1,$$

$$\text{with the definition } \text{Rect}\left(\frac{x}{a}, \frac{y}{b}\right) = \begin{cases} 1 & \text{if } (-a/2 \leq x \leq +a/2) \text{ and } (-b/2 \leq y \leq +b/2) \\ 0 & \text{otherwise} \end{cases}$$

$h^{(2)}(x, y)$ is then transformed into its hologram. Thus, the partly binarized hologram $H^{(1)}(\mu, \nu)$ is changed by the correction (5) into $H^{(2)}(\mu, \nu)$:

$$\begin{aligned} H^{(2)}(\mu, \nu) &= \mathcal{N}^{(2)}(\mathcal{R}e[\mathcal{F}\mathcal{T}\{h^{(2)}(x, y)\}] - Bs^{(2)}) \\ &= \mathcal{N}^{(2)}(H^{(1)}(\mu, \nu) + E^{(1)}(\mu, \nu) \\ &\quad - \left[\frac{2M_x M_y}{N^2} \text{sinc}\left(\frac{\mu M_x}{N}\right) \text{sinc}\left(\frac{\nu M_y}{N}\right) \cos\left(\frac{2\pi(x_0 \mu + y_0 \nu)}{N}\right) \right] \otimes E^{(1)}(\mu, \nu) - Bs^{(2)}) \end{aligned} \quad (6)$$

$$\mu, \nu = 0, 1, 2, \dots, N-1,$$

where \mathcal{FT} is the Fourier Transform, $\Re[\cdot]$ takes the real part of a complex function, $Bs^{(2)}$ is a constant bias, and $\mathcal{N}^{(2)}$ is normalization factor. $Bs^{(2)}$ and $\mathcal{N}^{(2)}$ are defined same as in Eq. (1). Thus, it is ensured that $0 \leq H^{(2)}(\mu, \nu) \leq 1$.

The bias term $Bs^{(2)}$ and the normalization factor $\mathcal{N}^{(2)}$ in $H^{(2)}(\mu, \nu)$ of (6) do not have any effects on the reconstruction quality measured by (3), since $Bs^{(2)}$ produces a delta function in the center of the reconstruction, which is outside of the reconstruction region w , and $\mathcal{N}^{(2)}$ is only an amplification. Therefore, $H^{(2)}(\mu, \nu)$ will still produce an error-free reconstruction in region w .

The difference between $H^{(2)}(\mu, \nu)$ and $H^{(1)}(\mu, \nu)$ is completely determined by the binarization error $E^{(1)}(\mu, \nu)$ of the K pixels. How these errors influence the hologram $H^{(2)}(\mu, \nu)$ at position (μ, ν) can be investigated by defining a generalized error diffusion coefficient $D^{(1)}_{\mu\mu'\nu\nu'}$:

$$H^{(2)}(\mu, \nu) = \mathcal{N}^{(2)} \left[H^{(1)}(\mu, \nu) + \sum_{(\mu', \nu')=(\mu_1, \nu_1)}^{(\mu_K, \nu_K)} D^{(1)}_{\mu\mu'\nu\nu'} E^{(1)}(\mu', \nu') - Bs^{(2)} \right] \quad (7)$$

$$\mu, \nu = 0, 1, 2, \dots, N-1.$$

That means by diffusing the binarization error $E^{(1)}(\mu, \nu)$ of Eq. (4) to the neighbors with diffusion coefficients $D^{(1)}_{\mu\mu'\nu\nu'}$, an error-free reconstruction in region w is achieved. From (6) and (7) above it follows:

$$\begin{aligned} & \sum_{(\mu', \nu')=(\mu_1, \nu_1)}^{(\mu_K, \nu_K)} D^{(1)}_{\mu\mu'\nu\nu'} E^{(1)}(\mu', \nu') \\ &= E^{(1)}(\mu, \nu) - \left[\frac{2M_x M_y}{N^2} \text{sinc}\left(\frac{\mu M_x}{N}\right) \text{sinc}\left(\frac{\nu M_y}{N}\right) \cos\left(\frac{2\pi(x_0 \mu + y_0 \nu)}{N}\right) \right] \otimes E^{(1)}(\mu, \nu) \\ &= \sum_{(\mu', \nu')=(\mu_1, \nu_1)}^{(\mu_K, \nu_K)} \delta_{\mu\mu'\nu\nu'} E^{(1)}(\mu', \nu') - \frac{2M_x M_y}{N^2} \sum_{(\mu', \nu')=(\mu_1, \nu_1)}^{(\mu_K, \nu_K)} \text{sinc}\left(\frac{(\mu - \mu') M_x}{N}\right) \\ & \quad \times \text{sinc}\left(\frac{(\nu - \nu') M_y}{N}\right) \cos\left(\frac{2\pi[x_0(\mu - \mu') + y_0(\nu - \nu')]}{N}\right) E^{(1)}(\mu', \nu') \\ &= \sum_{(\mu', \nu')=(\mu_1, \nu_1)}^{(\mu_K, \nu_K)} \left[\delta_{\mu\mu'\nu\nu'} - \frac{2M_x M_y}{N^2} \text{sinc}\left(\frac{(\mu - \mu') M_x}{N}\right) \text{sinc}\left(\frac{(\nu - \nu') M_y}{N}\right) \right. \\ & \quad \left. \times \cos\left(\frac{2\pi[x_0(\mu - \mu') + y_0(\nu - \nu')]}{N}\right) \right] E^{(1)}(\mu', \nu') \end{aligned}$$

where $\delta_{\mu\mu'\nu\nu'} = 1$ if $\mu = \mu'$ and $\nu = \nu'$, and $\delta_{\mu\mu'\nu\nu'} = 0$ otherwise.

It follows :

$$\begin{aligned} D^{(1)}_{\mu\mu'\nu\nu'} &= \delta_{\mu\mu'\nu\nu'} - \frac{2M_x M_y}{N^2} \text{sinc}\left(\frac{(\mu - \mu') M_x}{N}\right) \text{sinc}\left(\frac{(\nu - \nu') M_y}{N}\right) \cos\left(\frac{2\pi[x_0(\mu - \mu') + y_0(\nu - \nu')]}{N}\right) \\ &= \begin{cases} 1 - \frac{2M_x M_y}{N^2} & \text{if } \mu = \mu' \text{ and } \nu = \nu' \\ -\frac{2M_x M_y}{N^2} \text{sinc}\left(\frac{M_x(\mu - \mu')}{N}\right) \text{sinc}\left(\frac{M_y(\nu - \nu')}{N}\right) \cos\left(\frac{2\pi[x_0(\mu - \mu') + y_0(\nu - \nu')]}{N}\right) & \text{otherwise} \end{cases} \quad (8) \\ & \mu, \nu = 0, 1, 2, \dots, N-1; \quad (\mu', \nu') = (\mu_1, \nu_1), (\mu_2, \nu_2), \dots, (\mu_K, \nu_K). \end{aligned}$$

It shows that the generalized diffusion coefficients $D^{(1)}_{\mu\mu'\nu\nu'}$ depend only on the parameters M_x/N , M_y/N , x_0/N , y_0/N , on $(\mu - \mu')$ and $(\nu - \nu')$, they are independent of the step number of the

iteration, independent of which and how many pixels are binarized. In the following $D^{(1)}_{\mu\mu'v'v}$ is therefore denoted by $D(\mu-\mu',v-v')$.

Observing Eq. (7), we find that $\sum_{(\mu',v')} D(\mu-\mu',v-v')E^{(1)}(\mu',v')$ can be interpreted as a general version of filters in context of signal processing. The filter for the generation of binary holograms to produce an error-free reconstruction in region w is completely determined by $D(\mu-\mu',v-v')$. Since $D(\mu-\mu',v-v')$ produces zero-value in the reconstruction region, the sum $\sum_{(\mu',v')} D(\mu-\mu',v-v')E^{(1)}(\mu',v')$, which represents the convolution of $D(\mu,v)$ and $E^{(1)}(\mu,v)$, has zero-value too in this region. From Eq. (7), it follows immediately $h^{(2)}(x,y)=\mathcal{R}^{(2)}h^{(1)}(x,y)$ only in the reconstruction region, the surrounding area is filled with noise. Consequently, the desired object contained in $h^{(1)}(x,y)$ in region w is reproduced by $H^{(2)}(\mu,v)$.

The analysis above indicates if the binarization error $E^{(1)}(\mu,v)$ of the hologram is diffused to the neighbors with diffusion coefficients of Eq. (8), an error-free reconstruction in region w ($MSE=0$) is achieved. For this reason, we call $D(\mu-\mu',v-v')$ optimal diffusion coefficients.

Eq. (7) is an iterative representation of the binarization process. Nevertheless, this algorithm does not converge for further iterations. That is, MSE does not decrease for subsequent binarization of K pixels chosen randomly in $H^{(i)}(\mu,v)$ ($i \geq 2$). We say the binarization process stagnates. To find the reason of the stagnation, let us observe how the term $\sum_{(\mu',v')} D(\mu-\mu',v-v')E^{(1)}(\mu',v')$ in (7) influences the distribution of $H^{(1)}(\mu,v)$.

Normally the hologram distribution is locally correlated due to the carrier wave (reference wave). The correlation degree and direction are determined by the extension and fine structure of the signal, and by the direction of the carrier wave. The extension of the signal is just reflected in $D(\mu-\mu',v-v')$ by the size of the reconstruction region, and the direction of the carrier wave by the position where the reconstruction region is located, whereas the fine structure of the signal is reflected by the distribution of $E^{(1)}(\mu,v)$. One should pay attention that $D(\mu-\mu',v-v')$ has a minus sign except binarized pixels. That means, the binarization of K hologram pixels is done under penalty of smoothing other not binarized pixels. In other words, $\sum_{(\mu',v')} D(\mu-\mu',v-v')E^{(1)}(\mu',v')$ smears the distribution of $H^{(1)}(\mu,v)$ except K binarized pixels. The following two cases are considered:

Case one: the number of binarized pixels K is very small compared to N^2 . Thus, K pixels chosen in the $(i+1)$ -th iteration ($i=1,2,\dots$) fall into the effective diffusion area of the K binarized pixels of the i -th iteration with small probability, because they are randomly chosen, independent of each other. Here the effective diffusion area means that two pixels just overlap, or are adjacent. Consequently, the distribution of binarization error of the $(i+1)$ -th iteration is similar to that of the i -th iteration statistically. It is difficult to predict the reconstruction quality in region w , in comparison with that of the i -th iteration. In other words, $MSE^{(i+1)}$ and $MSE^{(i)}$ do not have a clear relation in this case. One can not say the algorithm converge or not.

Case two: K is large enough but smaller than N^2 . In contrast to case one, there are many pixels chosen in the $(i+1)$ -th iteration falling into the effective diffusion area of the K binarized pixels of the i -th iteration. Obviously, if a pixel binarized in the $(i+1)$ -th iteration is the one that was binarized in the i -th iteration, $E^{(i+1)}(\mu,v) \leq E^{(i)}(\mu,v)$ is resulted. We say it results in a constructive reconstruction in region w . If a pixel binarized in the $(i+1)$ -th iteration is the one that was smoothed in the i -th iteration (neighborhood of a binarized pixel), $E^{(i+1)}(\mu,v) \geq E^{(i)}(\mu,v)$ happens with high probability. We say it results in a destructive reconstruction in region w . Moreover, many pixels having values near the threshold value T in the i -th iteration have changed their relationship to T due to the smoothness, i.e., $H^{(i)}(\mu,v) \geq T$ to $H^{(i+1)}(\mu,v) < T$, or $H^{(i)}(\mu,v) < T$ to $H^{(i+1)}(\mu,v) \geq T$. Thus, the binarization error $E^{(i+1)}(\mu,v)$ of such pixels changes sign compared to $E^{(i)}(\mu,v)$, which seriously destroys the reconstruction in region w . As a result, when K is sufficiently large, destructive reconstruction is stronger than constructive in the $(i+1)$ -th iteration statistically. It results in $MSE^{(i+1)} \geq MSE^{(i)}$, or the binarization process stagnates.

The analysis above indicates the stagnation can be avoided if the smoothing to the hologram distribution (see (7)) is slowed down, so that the constructive reconstruction is dominate

statistically. This can be realized by enhancing the signal in the reconstruction region w . We modify the correction of (5) into:

$$h^{(i+1)}(x,y)=h^{(i)}(x,y)+e^{(i)}(x,y)-\left[\text{Rect}\left(\frac{x-x_0}{M_x},\frac{y-y_0}{M_y}\right)+\text{Rect}\left(\frac{x+x_0}{M_x},\frac{y+y_0}{M_y}\right)\right]e^{(i)}(x,y) \\ +\beta^{(i)}\left[\text{Rect}\left(\frac{x-x_0}{M_x},\frac{y-y_0}{M_y}\right)+\text{Rect}\left(\frac{x+x_0}{M_x},\frac{y+y_0}{M_y}\right)\right]h^{(i)}(x,y) \quad (9)$$

$x, y=0, 1, 2, \dots, N-1,$

where $\beta^{(i)} \geq 0$ represents the enhancement of the signal in region w . We call $\beta^{(i)}$ enhancement factor.

Then, (9) is transformed into its hologram:

$$H^{(i+1)}(\mu, \nu) = \mathcal{N}^{(i+1)} \left[H^{(i)}(\mu, \nu) + \frac{1}{1+\beta^{(i)}} \sum_{(\mu', \nu')=(\mu_1, \nu_1)}^{(\mu_K, \nu_K)} D(\mu-\mu', \nu-\nu') E^{(i)}(\mu', \nu') \right. \\ \left. - \frac{\beta^{(i)}}{1+\beta^{(i)}} \sum_{(\mu', \nu')=(\mu_1, \nu_1)}^{(\mu_K, \nu_K)} D(\mu-\mu', \nu-\nu') H^{(i)}(\mu', \nu') - B_s^{(i+1)} \right] \quad (10)$$

$\mu, \nu=0, 1, 2, \dots, N-1.$

Eq. (10) represents a general binarization process. We call $H^{(i+1)}(\mu, \nu)$ pseudo-binary hologram for $i \geq 1$. Note that pseudo-binary holograms produce an error-free reconstruction only in region w .

Two terms in (10) play the role to resist the smoothing of the hologram distribution $H^{(i)}(\mu, \nu)$. $\sum_{(\mu', \nu')} D(\mu-\mu', \nu-\nu') E^{(i)}(\mu', \nu')$ is reduced by factor $(1+\beta^{(i)})$ compared to (7). From the analyses above, the smoothing is reduced. The other term $-\beta^{(i)}/(1+\beta^{(i)}) \sum_{(\mu', \nu')} D(\mu-\mu', \nu-\nu') H^{(i)}(\mu', \nu')$ even enlarges the contrast of the hologram, under penalty of reducing the binarized pixel value.

If $\beta^{(i)}$ is too small, (10) degenerates into (7), the binarization process stagnates as analyzed above. On the other hand, if $\beta^{(i)}$ is too large, that is, the signal is over enhanced, the binarization process stagnates too. In this case, (10) becomes

$$H^{(i+1)}(\mu, \nu) = \mathcal{N}^{(i+1)} \left[H^{(i)}(\mu, \nu) - \sum_{(\mu', \nu')=(\mu_1, \nu_1)}^{(\mu_K, \nu_K)} D(\mu-\mu', \nu-\nu') H^{(i)}(\mu', \nu') - B_s^{(i+1)} \right] \\ = \mathcal{N}^{(i+1)} \left[\frac{2M_x M_y}{N^2} \text{sinc}\left(\frac{\mu M_x}{N}\right) \text{sinc}\left(\frac{\nu M_y}{N}\right) \cos\left(\frac{2\pi(x_0 \mu + y_0 \nu)}{N}\right) \otimes H^{(i)}(\mu, \nu) - B_s^{(i+1)} \right]$$

which does not produce any binarization noise inside and outside the reconstruction region. Over enhancement of the signal has the effect that the binarization noise in the entire reconstruction plane is suppressed after normalization. This is much similar to the reconstruction of $H^{(1)}(\mu, \nu)$. Consequently, $H^{(i+1)}(\mu, \nu) \rightarrow H^{(1)}(\mu, \nu)$, the binarization process is slowed down, or in extreme case stagnates.

These show the importance and subtlety of the enhancement factor $\beta^{(i)}$ to the convergence of the algorithm, so it must be determined rationally.

Actually, the basic idea of iterative algorithms comes from the iterative Fourier transform algorithm [17]. The iterative binarization algorithm described above uses the same principle, but different constraints. The significance of Eq. (10) is that the connections between hologram domain and reconstruction domain are cut off. Consequently, the generation of a binary hologram from its gray-scale one becomes repeatedly performing error diffusion processes. In section 6, two approaches are suggested to determine $\beta^{(i)}$. The error in the reconstruction region can also be estimated only in hologram domain (see section 4).

The flow-chart shown in Fig. 1 depicts the iterative binarization process, where the stop criterion can be set as the reduction degree of $MSE^{(i)}$, for example, or the number of iterations.

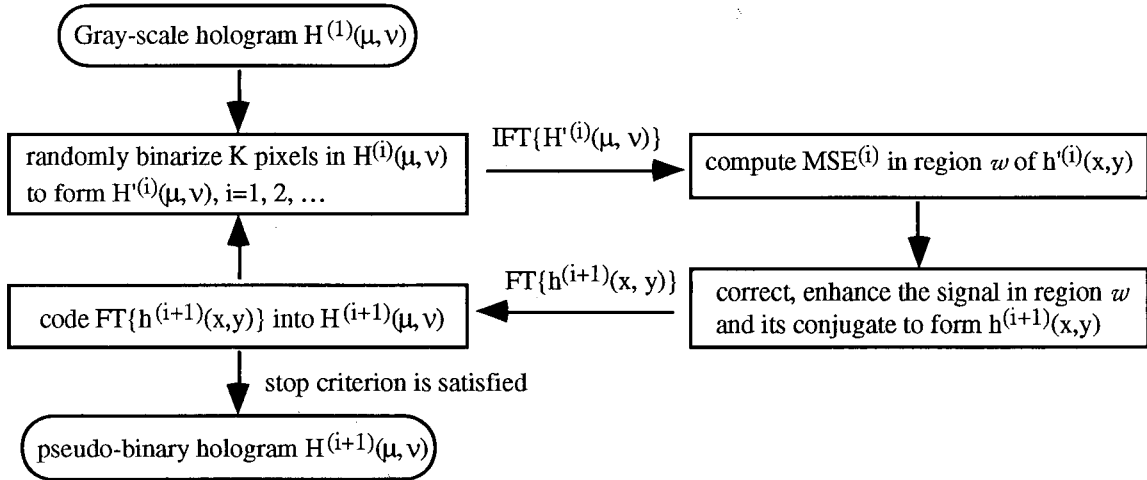


Fig. 1. Flow-chart of iterative binarization algorithm.

Normally $D(0,0) < 1$. Thus, pseudo-binary holograms are still gray-scale holograms, but always tending to binary ones. Therefore, it is theoretically impossible to generate a binary hologram with that to produce an error-free reconstruction in region w . Even a partly binarized hologram ($K < N^2$) cannot be expected to produce the same reconstruction quality in region w as a gray-scale hologram does. We conclude that binarization noise in region w can only be reduced, but not completely eliminated.

In next sections we investigate under which conditions the reconstruction can be improved, that is, under which conditions MSE can be reduced. For noniterative algorithms we use Eq. (7), whereas for iterative algorithms we use Eq. (10) and Fig. 1.

3. Properties of diffusion coefficients $D(\mu-\mu', \nu-\nu')$

The procedure performed in the preceding section for an error-free reconstruction in region w requires region w , its conjugate, and the delta function in the reconstruction do not overlap each other. That means, parameters M_x , M_y , N , x_0 and y_0 in $D(\mu-\mu', \nu-\nu')$ must be restricted in allowed value ranges. We first determine these ranges and their implicit relations before discussing $D(\mu-\mu', \nu-\nu')$.

- $(-N/2, -N/2) \leq (x_0, y_0) < 0$, or $0 < (x_0, y_0) < (+N/2, +N/2)$.
- $0 < M_x \leq N$, $0 < M_y \leq N$.
- At least, one of the following conditions must be fulfilled:
 $(M_x < N/2)$ and $(M_x/2 < |x_0| \leq (N-M_x)/2)$,
 $(M_y < N/2)$ and $(M_y/2 < |y_0| \leq (N-M_y)/2)$.

Taking into account of these restrictions, some properties of the diffusion coefficients $D(\mu-\mu', \nu-\nu')$ are listed below. The binarized pixel is denoted by (μ', ν') .

- $D(\mu-\mu', \nu-\nu')$ decrease oscillatedly with an increment of $|\mu-\mu'|$ and $|\nu-\nu'|$; (11a)
- $D(\mu-\mu', \nu-\nu')$ is symmetric to pixel (μ', ν') ; (11b)
- $D(\mu-\mu', \nu-\nu') \rightarrow \delta_{\mu\mu', \nu\nu'}$ if $M_x \rightarrow 0$ and $M_y \rightarrow 0$; (11c)
- $D(\mu-\mu', \nu-\nu') \rightarrow 0$ if $M_x M_y \rightarrow N^2/2$; (11d)
- $\sum_{(\mu, \nu)} D(\mu-\mu', \nu-\nu') \rightarrow 1^1$. (11e)

1)

This result is obtained by using the following integral:

Two examples shown in Fig. 2 illustrate the behavior of coefficients $D(\mu-\mu',v-v')$ in a 7×7 neighborhood around binarized pixel (μ',v') , where x_0 and y_0 remain unchanged. In the first example, Fig. 2(a), $M_x=M_y=3N/8$, $(x_0,y_0)=(N/4,N/4)$; and in the second example, Fig. 2(b), $M_x=M_y=N/8$, $(x_0,y_0)=(N/4,N/4)$ too. Numerical values of $D(-',v-v')$ corresponding to Fig. 2(a) and Fig. 2(b) are given in Table I(a) and I(b) respectively.

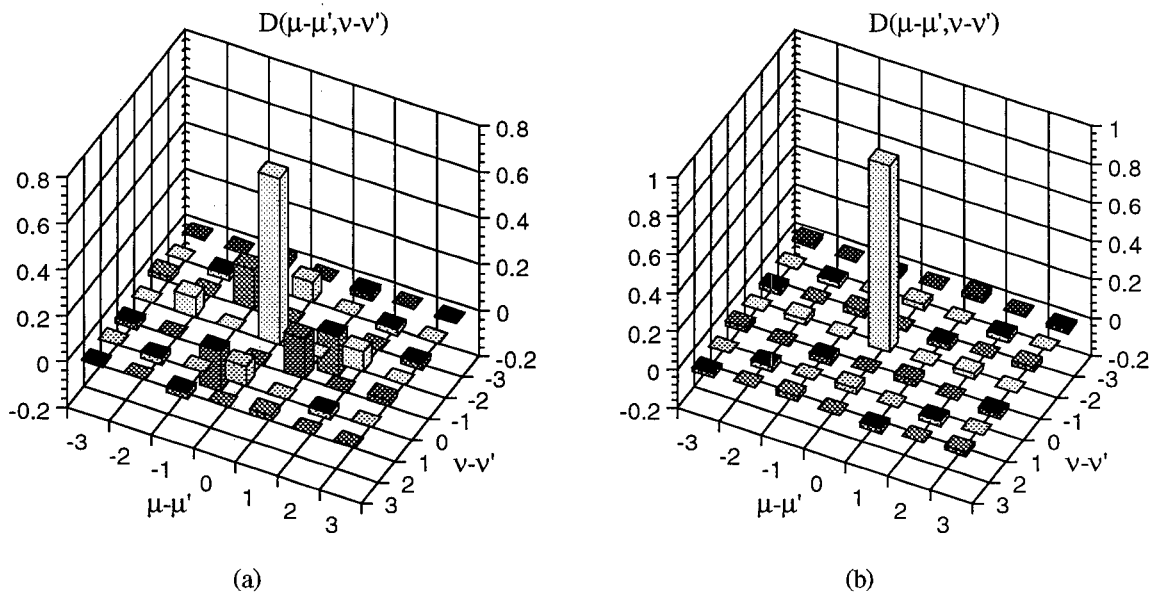


Fig. 2. Dependence of diffusion coefficients $D(-',v-v')$ on parameters M_x and M_y , where $(x_0,y_0)=(N/4,N/4)$. (a) $M_x=M_y=3N/8$, (b) $M_x=M_y=N/8$. Black posts represent negative values.

Table I. Values of $D(-',v-v')$ corresponding to Fig. 2(a) for $M_x=M_y=3N/8$ and Fig. 2(b) for $M_x=M_y=N/8$. $(x_0,y_0)=(N/4,N/4)$.

(a) $(M_x=M_y=3N/8) \rightarrow$

\downarrow	+0.00330	0	+0.02388	0	-0.02388	0	-0.00330
v	0	-0.02533	0	+0.08440	0	-0.02533	0
	+0.02388	0	+0.17297	0	-0.17297	0	-0.02388
	0	+0.08440	0	$+23/32 \left(\frac{v}{N} \right)$	0	+0.08440	0
	-0.02388	0	-0.17297	0	+0.17297	0	+0.02388
	0	-0.02533	0	+0.08440	0	-0.02533	0
	-0.00330	0	-0.02388	0	+0.02388	0	+0.00330

(b) $(M_x=M_y=N/8) \rightarrow$

$$\int\limits_{-\infty}^{+\infty} \text{sinc}(ax) \cos(2\pi bx) \, dx = \begin{cases} \frac{1}{a} & \text{for } \left| \frac{2b}{a} \right| < 1 \\ \frac{1}{2a} & \text{for } \left| \frac{2b}{a} \right| = 1 \\ 0 & \text{for } \left| \frac{2b}{a} \right| > 1 \end{cases}$$

↓	+0.01922	0	-0.02388	0	+0.02388	0	-0.01922
v	0	-0.02533	0	+0.02813	0	-0.02533	0
	-0.02388	0	+0.02968	0	-0.02968	0	+0.02388
	0	+0.02813	0	+31/32 (\downarrow, \downarrow)	0	+0.02813	0
	+0.02388	0	-0.02968	0	+0.02968	0	-0.02388
	0	-0.02533	0	+0.02813	0	-0.02533	0
	-0.01922	0	+0.02388	0	-0.02388	0	+0.01922

The binarized pixel is denoted by (\downarrow, \downarrow) in Table I. It has been shown that the largest error component is diffused to the binarized pixel itself (23/32 and 31/32 for the examples above respectively). With a reduction of the reconstruction region (M_x and M_y change from $3N/8$ to $N/8$), a larger error amount (change from 23/32 to 31/32) is diffused to the binarized pixel, whereas errors diffused to its neighbors change from a small region and a large amount to a large region and a small amount. In other words, the smaller the reconstruction region, the more error is diffused to the binarized pixel and the less to its neighbors.

Two more examples shown in Fig. 3 illustrate the dependence of diffusion coefficients $D(\downarrow, \downarrow - v - v')$ on the position of the reconstruction region (x_0, y_0) , where M_x and M_y remain unchanged. Related values to Fig. 3 are given in Table II.

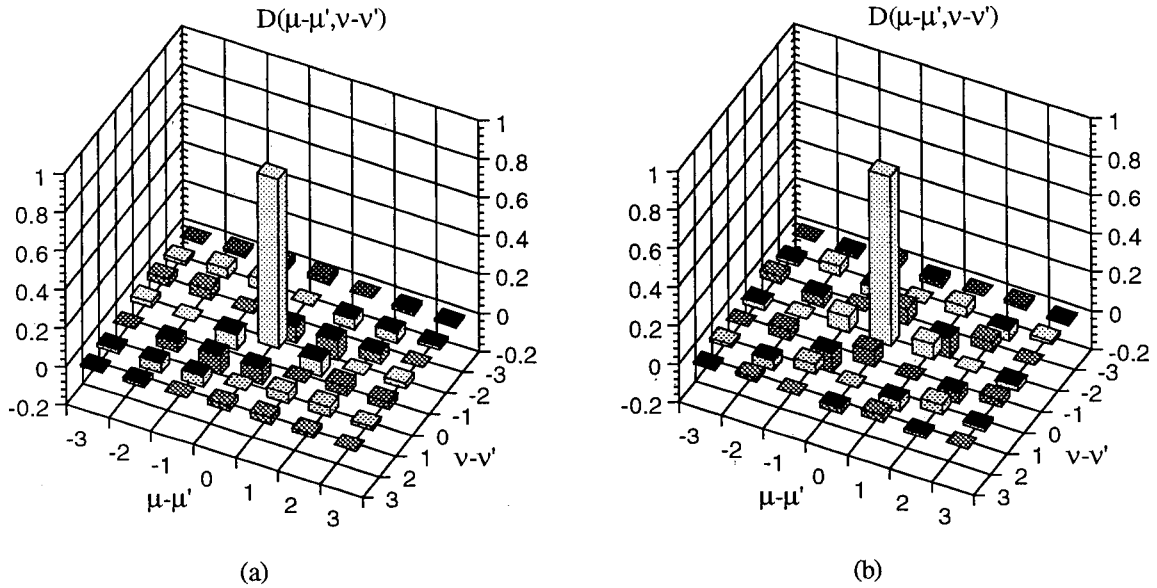


Fig. 3. Dependence of diffusion coefficients $D(\downarrow, \downarrow - v - v')$ on parameter (x_0, y_0) , where $M_x = M_y = N/4$. (a) $(x_0, y_0) = (N/8, N/8)$, (b) $(x_0, y_0) = (3N/8, 3N/8)$. Black posts represent negative values.

Table II. Values of $D(\downarrow, \downarrow - v - v')$ corresponding to Fig. 3(a) for $(x_0, y_0) = (N/8, N/8)$ and Fig. 3(b) for $(x_0, y_0) = (3N/8, 3N/8)$. $M_x = M_y = N/4$.

(a) $(x_0, y_0) = (N/8, N/8) \rightarrow$

↓	0	+0.01689	+0.03377	+0.02653	0	-0.01689	-0.01126
v	+0.01689	+0.05066	+0.05066	0	-0.05066	-0.05066	-0.01689
	+0.03377	+0.05066	0	-0.07958	-0.10132	-0.05066	0
	+0.02653	0	-0.07958	+7/8 (\downarrow, \downarrow)	-0.07958	0	+0.02653

0	-0.05066	-0.10132	-0.07958	0	+0.05066	+0.03377
-0.01689	-0.05066	-0.05066	0	+0.05066	+0.05066	+0.01689
-0.01126	-0.01689	0	+0.02653	+0.03377	+0.01689	0

(b) $(x_0, y_0) = (3N/8, 3N/8) \rightarrow$

↓	0	-0.01689	+0.03377	-0.02653	0	+0.01689	-0.01126
v	-0.01689	+0.05066	-0.05066	0	+0.05066	-0.05066	+0.01689
	+0.03377	-0.05066	0	+0.07958	-0.10132	+0.05066	0
	-0.02653	0	+0.07958	+7/8 ('v')	+0.07958	0	-0.02653
	0	+0.05066	-0.10132	+0.07958	0	-0.05066	+0.03377
	+0.01689	-0.05066	+0.05066	0	-0.05066	+0.05066	-0.01689
	-0.01126	+0.01689	0	-0.02653	+0.03377	-0.01689	0

It has been shown that for different (x_0, y_0) (here $(x_0, y_0) = (N/8, N/8), (3N/8, 3N/8)$) the same error amount is diffused to the binarized pixel itself ('v'), and the same error amounts but with different signs are diffused to its neighbors. Here we have only considered the case $(x_0, y_0) > (0, 0)$. For other cases the results are similar due to the property of the cosine function in $D(-', v-v')$. With an enlargement of $|x_0|$ and $|y_0|$, $D(-', v-v')$ changes its signs more frequently, that is, the modulation from the cosine function becomes much stronger (in the main diagonal direction, $D(-', v-v')$ changes its sign 3 times and 9 times for our two examples above, respectively). These results will be used to estimate the reconstruction quality in region w .

4. Error estimation

In this section we investigate how the diffusion coefficients $D(-', v-v')$ and binarization error $E^{(i)}(, v)$ influence the reconstruction quality in region w . First, we discuss the influence of parameters in $D(-', v-v')$. Two extreme cases are considered:

Case one: $M_x \rightarrow 0$ and $M_y \rightarrow 0$, i.e., the reconstruction region is very small. In this case we have $D(-', v-v') \rightarrow \delta_{, w}$ according to (11c) of section 3. This means the binarization error is principally diffused to the binarized pixel, hardly to its neighbors. From (7), it follows $H^{(2)}(, v) \rightarrow \mathcal{H}^{(2)}[H^{(1)}(, v) + E^{(1)}(, v) - B_s^{(2)}]$. That is, $H^{(2)}(, v)$ is much near to a binary hologram. The binarization does not introduce serious noise in the reconstruction region. From (10), $H^{(i+1)}(, v) \rightarrow \mathcal{H}^{(i+1)}[H^{(i)}(, v) + E^{(i)}(, v) - B_s^{(i+1)}]$, i.e., (10) degenerates into Eq. (7). As a result, noniterative algorithms achieve the same reconstruction quality as iterative algorithms do, with $MSE \rightarrow 0$.

Case two: $M_x \rightarrow N/2$ and $M_y \rightarrow N$ (or $M_x \rightarrow N$ and $M_y \rightarrow N/2$), i.e., the reconstruction region has the size nearly the half of the entire reconstruction plane. In this case $D(-', v-v') \rightarrow 0$ according to (11d) of section 3. This means the binarization error can be diffused neither to the binarized pixel, nor to its neighbors. From (7), $H^{(2)}(, v) \rightarrow H^{(1)}(, v)$. That is, one can not generate a pseudo-binary hologram to achieve an improved reconstruction. For iterative binarization process, it leads to $H^{(i+1)}(, v) \rightarrow H^{(i)}(, v)$ (see (10)), $MSE^{(i)} \sim MSE^{(1)}$ statistically, and the binarization process stagnates.

These results are well understood in the reconstruction domain. In case one, since the reconstruction region is small, the surrounding area is large. Thus, there is more free space available to distribute binarization noise into there. As a result, the reconstruction in region w can significantly be improved. In contrast, there is no space available in case two to distribute binarization noise. Consequently, the reconstruction in region w can not be improved at all.

Parameters x_0 and y_0 in $D(-', v-v')$ have no influence on the reconstruction quality.

In the following we investigate the influence of $E^{(i)}(x, y)$ on the reconstruction quality. Using the relation $h^{(i)}(x, y) = h^{(i)}(x, y) + e^{(i)}(x, y)$, we can rewrite the formula of MSE in (3) in form:

$$MSE^{(i)} = \frac{\sum_w |e^{(i)}(x, y)|^2 - |\sum_w h^{(i)}(x, y) e^{(i)*}(x, y)|^2 / \sum_w |h^{(i)}(x, y)|^2}{\sum_w |e^{(i)}(x, y)|^2 + \sum_w |h^{(i)}(x, y)|^2 + 2\text{Re}[\sum_w h^{(i)}(x, y) e^{(i)*}(x, y)]} \quad (12)$$

which is represented by the reconstruction $h^{(i)}(x, y)$ of the pseudo-binary hologram, and noise $e^{(i)}(x, y)$ caused by the binarization of the pseudo-binary hologram.

Mapping all terms in (12) into hologram domain, i.e., $H_w |e^{(i)}(x, y)|^2$, $H_w |h^{(i)}(x, y)|^2$ and $H_w h^{(i)}(x, y) e^{(i)*}(x, y)$, a direct relation between binarization error $E^{(i)}(x, y)$ in the hologram and resulting MSE in the reconstruction region can be established. For example,

$$\begin{aligned} H_w h(x, y) e^*(x, y) &= \sum_w h(x, y) \frac{1}{N} \sum_{\mu, \nu} E(\mu, \nu) \exp\left\{\frac{-j2\pi(\mu x + \nu y)}{N}\right\} \\ &= \sum_{\mu, \nu} E(\mu, \nu) \frac{1}{N} \sum_w h(x, y) \exp\left\{\frac{-j2\pi(\mu x + \nu y)}{N}\right\} \\ &= \sum_{\mu, \nu} E(\mu, \nu) \frac{1}{N} \sum_{x, y} \text{Rect}\left(\frac{x-x_0}{M_x}, \frac{y-y_0}{M_y}\right) h(x, y) \exp\left\{\frac{-j2\pi(\mu x + \nu y)}{N}\right\} \\ &= \sum_{\mu, \nu} E(\mu, \nu) \left[\frac{M_x M_y}{N^2} \text{sinc}\left(\frac{M_x \mu}{N}\right) \text{sinc}\left(\frac{M_y \nu}{N}\right) \exp\left(\frac{-j2\pi(x_0 \mu + y_0 \nu)}{N}\right) \otimes H(\mu, \nu) \right] \\ &= H_{\mu, \nu} E(\mu, \nu) H_{\mu, \nu} C(-\mu, -\nu) H(\mu, \nu) \quad (13a) \end{aligned}$$

$$= H_{\mu, \nu} C(-\mu, -\nu) [H(\mu, \nu) * E(\mu, \nu)] \quad (13b)$$

$$\text{where } C(-\mu, -\nu) = \frac{M_x M_y}{N^2} \text{sinc}\left(\frac{M_x(\mu - \mu')}{N}\right) \text{sinc}\left(\frac{M_y(\nu - \nu')}{N}\right) \exp\left\{\frac{-j2\pi[x_0(\mu - \mu') + y_0(\nu - \nu')]}{N}\right\},$$

which we call coupling coefficient; * denotes the correlation.

Similarly, we have

$$H_w |e(x, y)|^2 = H_w e(x, y) e^*(x, y) = H_{\mu, \nu} C(-\mu, -\nu) [E(\mu, \nu) * E(\mu, \nu)] \quad (13c)$$

$$H_w |h(x, y)|^2 = H_w h(x, y) h^*(x, y) = H_{\mu, \nu} C(-\mu, -\nu) [H(\mu, \nu) * H(\mu, \nu)] \quad (13d)$$

To simplify the notation we define

$$ACB = H_{\mu, \nu} A(\mu, \nu) H_{\mu, \nu} C(-\mu, -\nu) B(\mu, \nu) = H_{\mu, \nu} C(-\mu, -\nu) [A(\mu, \nu) * B(\mu, \nu)].$$

Thus, MSE is evaluated in hologram domain by

$$MSE^{(i)} = \frac{E^{(i)} C E^{(i)} - |H^{(i)} C E^{(i)}|^2 / H^{(i)} C H^{(i)}}{E^{(i)} C E^{(i)} + H^{(i)} C H^{(i)} + H^{(i)} (C^* + C) E^{(i)}} \quad (14)$$

Coupling coefficient $C(-\mu, -\nu)$ is related to diffusion coefficients $D(-\mu, -\nu)$ by

$$D(-\mu, -\nu) = \delta_{\mu, \nu} - [C^*(-\mu, -\nu) + C(-\mu, -\nu)]$$

Similar to $D(-\mu, -\nu)$ in (11), we have:

$$C(-\mu, -\nu) \text{ decrease in magnitude oscillatedly with an increment of } 1/N \text{ and } |\nu - \nu'|; \quad (15a)$$

$$C(-\mu, -\nu) = C^*[-(-\mu), -(-\nu)]; \quad (15b)$$

$$C(-\mu, -\nu) \rightarrow 0 \text{ (constant) if } M_x \rightarrow 0 \text{ and } M_y \rightarrow 0; \quad (15c)$$

$$[C(-\mu, -\nu) + C^*(-\mu, -\nu)] \rightarrow \delta_{\mu, \nu} \text{ if } M_x M_y \rightarrow N^2/2; \quad (15d)$$

$$H_{\mu, \nu} C(-\mu, -\nu) \rightarrow 0. \quad (15e)$$

The coupling coefficient $C(-', v - v')$ in MSE of (14) depends on parameters M_x/N , M_y/N , x_0/N and y_0/N too. That is, for a given application, coupling coefficient $C(-', v - v')$ is unchanged. Thus, from (14) MSE is tightly connected with integrals of auto-correlation of E , of H , and cross-correlation of E and H weighted by $C(-', v)$.

Different applications are possessed of different coupling degree. The example $M_x \rightarrow 0$ and $M_y \rightarrow 0$ discussed above leads to $C(-', v - v') \rightarrow 0$ (constant). We say it is strongly coupled. The example $M_x \rightarrow N/2$ and $M_y \rightarrow N$ (or $M_x \rightarrow N$ and $M_y \rightarrow N/2$) leads to $C(-', v - v') \rightarrow 1/2 \delta_{vv'} [1 - \exp\{-j(-')\}]/j(-')$ (or $C(-', v - v') \rightarrow 1/2 \delta_{-v-v'} [1 - \exp\{-j(v - v')\}]/j(v - v')$). We say it is decoupled in v -direction (or in $-$ -direction).

Eq. (14) gives a general convergence condition for iterative algorithms. The introduction of enhancement factor $\beta^{(i)}$ alters the distribution of the hologram $H^{(i+1)}(-', v)$ (see (10)), and therefore the distribution of binarization error $E^{(i+1)}(-', v)$. Thus, each term in (14) is changed. That is, $MSE^{(i+1)}$ depends not only on the error distribution of $E^{(i+1)}(-', v)$, but also on the distribution of $H^{(i+1)}(-', v)$, and correlation of $E^{(i+1)}(-', v)$ and $H^{(i+1)}(-', v)$. An appropriate $\beta^{(i)}$ can lead to $MSE^{(i+1)} \ll MSE^{(i)}$. However, it is quite sophisticated analytically to determine the value of $\beta^{(i)}$. In practice, one can try it for a given application (see section 6).

If all pixels in the hologram are binarized one time, i.e., the number of binarized pixels K is equal to N^2 without overlaps, the binarization process stagnates too [14]. This can be understood as follows: the binarization error diffused to the pixel $(-', v)$ ($-', v=0, 1, 2, \dots, N-1$) from its neighbors $(-, v)$ ($-, v=0, 1, 2, \dots, N-1$, $(-, v)(-, v')$) is always smaller than the binarization error from itself, i.e., $H(-', v), (-', v)(-, v) D(-', v - v') E^{(i)}(-', v) < D(0, 0) E^{(i)}(-', v)$ (see (10) and (11e)). Thus the binary state formed by the latter cannot be changed by the former. That means, the binary states formed by the 1-th binarization process cannot be changed by the 2-th or subsequent binarization processes. In other words, the hologram has arrived at a steady distribution state. In practice, we choose K sufficiently large, but smaller than N^2 .

5. Noniterative method

In the following we discuss a noniterative method: error diffusion [5-8]. This algorithm binarizes holograms sequentially. The number of binarized pixels increases by one each time until to N^2 .

Error diffusion was originally developed for displaying gray-value images on black and white screens by half-toning [4]. Hauck and Bryngdahl applied this technique to computer holography [5]. The basic concept can be described as follows: The first pixel of the hologram $H(-', v)$, $H(0, 0)$, is compared with a threshold T , and a new hologram value $H_{out}(0, 0)$ is produced. $H_{out}(0, 0) = 1$ if $H(0, 0) \geq T$, otherwise $H_{out}(0, 0) = 0$. This threshold operation introduces an error $E(0, 0) = H_{out}(0, 0) - H(0, 0)$, which is diffused to those pixels in the neighborhood of $(0, 0)$ that have not been processed yet. We use $(-, v')$ to denote the pixel being binarized, then the new value of an unprocessed pixel located at $(-, v)$, $(-, v) = (+1, v')$, $(+1, v'+1)$, $(-, v'+1)$, $(-1, v'+1)$, is $H'(-', v) = H(-', v) + W \cdot_{vv'} E(-', v')$, where $W \cdot_{vv'}$ are called diffusion coefficients. This algorithm is sequential, i.e., it works line by line from left to right.

Many different variations [7-10] of this algorithm have been suggested. They include the use of non-adjacent pixels, and of different diffusion coefficients which could also vary from pixel to pixel. Using the results of section 3, we can satisfactorily explain which variation should be used for which application. Below, we present this analysis and show the limits of the error diffusion algorithm.

Rather than choosing pixels randomly for binarization as described in section 2, they are chosen here one by one sequentially. The binarization of the first pixel $H(0, 0)$ has already introduced a binarization error, which has to be diffused to its neighbors. How to diffuse this error becomes the core of the algorithm. Some improvements have been done using variations of diffusion coefficients [8-10]. However, none of these variations has taken into account of the effect of the size of the reconstruction region. Therefore, their application range is limited. The mathematical analysis of section 2 and from it derived results have shown that the optimal diffusion coefficients and the diffusion size are closely related to the applications, i.e.,

they depend on the size of the reconstruction region, M_x/N and M_y/N , on the position where the reconstruction region is located, x_0/N and y_0/N (see (8)).

If the binarization error is diffused to the neighbors using the optimal diffusion coefficients $D(-', v-v')$, $MSE=0$ in region w is achieved. However, we could never generate a binary hologram. Using the ordinary error diffusion algorithm as described above, we can generate a binary hologram, but we can never achieve $MSE=0$, because the binarization error cannot be diffused to the pixels that have already been processed; besides, the binarized pixel itself always introduces error provided $M_x > 0$ and $M_y > 0$ as indicated in (8). This is one of limitations of the error diffusion algorithm.

In the following, a combination of the ordinary error diffusion algorithm with the optimal diffusion coefficients is used to generate binary holograms. The diffusion coefficients are then:

$$D^{(ED)}(-', v-v') = \begin{cases} 1 & \text{if } (\mu=\mu') \text{ and } (v=v') \\ D(\mu-\mu', v-v') & \text{if } (v \geq v' \text{ and } \mu > \mu') \text{ or } (v > v' \text{ and } \mu \leq \mu') \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

and

$$H'(\cdot, v) = H(\cdot, v) + H'(\cdot, v) D^{(ED)}(-', v-v') E'(\cdot, v) \quad (17)$$

In addition, one should choose an appropriate diffusion size. The sum is to all processed pixels (\cdot, v') , or the predecessors of (\cdot, v) , within the diffusion size.

Asymmetrical diffusion coefficients of error diffusion algorithms introduce disturbances in amplitude, and principally in phase inside and outside the reconstruction region. Therefore, this algorithm is more suitable for applications where the reconstruction of intensity signals is required. We define the following Mean-squared Error for intensity signals:

$$MSE^\# = 1.0 - \frac{[\sum_w |u(x-x_0, y-y_0)| |h^*(x, y)|]^2}{\sum_w |u(x-x_0, y-y_0)|^2 \sum_w |h(x, y)|^2} \quad (18)$$

where $u(x, y)$ and $h(x, y)$ are defined same as in (3), $\#$ denotes the difference from MSE defined in (3) for complex signals.

The example below shows the error reduction degree and the efficiency of the algorithm (16) and (17).

The object to be reconstructed is the capital letter 'G' represented by a square region of 37×43 pixels. It is centrally embedded into a zero array of 128×128 pixels as shown in Fig. 4(a). Random phases are assigned to it to generate its gray-scale hologram. Fig. 4(b) shows this hologram with $x_0=y_0=32$, and 256 gray levels.

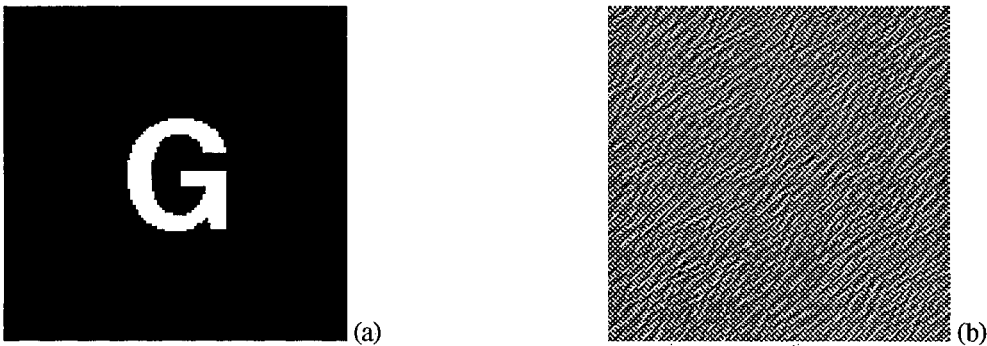


Fig. 4. Original object (capital letter 'G') (a), and its gray scale hologram (256 gray levels) (b).

The reconstruction region w has the size same as the object, namely $M_x \times M_y = 37 \times 43$ pixels. The center of the reconstruction region is $(x_0, y_0) = (32, 32)$. Fig. 5(a) shows the binary hologram of Fig. 4(b) generated using (16) and (17), and diffusion size of 9-adjacent neighborhoods in $-$ direction and 7-adjacent neighborhoods in v -direction. Fig. 5(b) shows

the computer simulated reconstruction of Fig. 5(a) with $MSE=0.1847$, $MSE^{\#}=0.0319$. The diffraction efficiency is given by $\eta=3.29\%$.

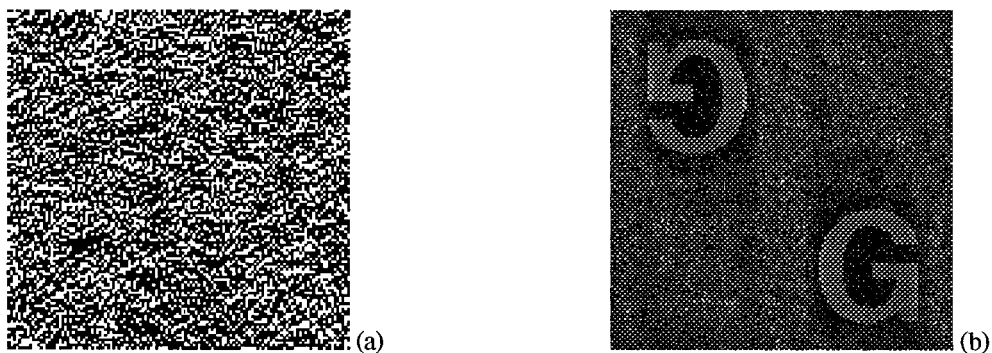


Fig. 5. Binary hologram generated using the error diffusion algorithm proposed here (Eqs. (16) and (17)) and diffusion size of 9-adjacent neighborhoods in x -direction and 7-adjacent neighborhoods in y -direction (a), and its simulated reconstruction (b).

For comparison, another binary hologram of Fig. 4(b) is generated using the error diffusion algorithm given by Ref. [9]. The basic idea of this algorithm can be described as follows: one designs a filter for gray-scale holograms. This filter introduces a zero in the center of the reconstruction region. In addition, this filter function should remain the slope of the neighborhood of the zero point possibly small. The filter function is the diffusion coefficient matrix. In the reconstruction domain it is

$$h'(x,y)=h^{(1)}(x,y)+f(x,y)e(x,y)$$

The reconstruction $h'(x,y)$ of the binary hologram $H^{(1)}(x,y)$ is the reconstruction of the gray-scale hologram $H^{(1)}(x,y)$ plus a noise term $f(x,y)e(x,y)$. $f(x,y)$ is the transfer function of the filter, and is zero in the center of the reconstruction region.

As indicated in [9], large diffusion size reduces the filter effect. Besides, $f(x,y) \neq 0$ in the reconstruction region does not guarantee $f(x,y)e(x,y) \neq 0$, i.e., the algorithm is not always stable. We use the diffusion coefficients given by [9] to binarize the gray-scale hologram Fig. 4(b). Fig. 6(a) shows the generated binary hologram. Its simulated reconstruction is shown in Fig. 6(b) with $MSE=0.1515$, $MSE^{\#}=0.0946$, and $\eta=1.24\%$.

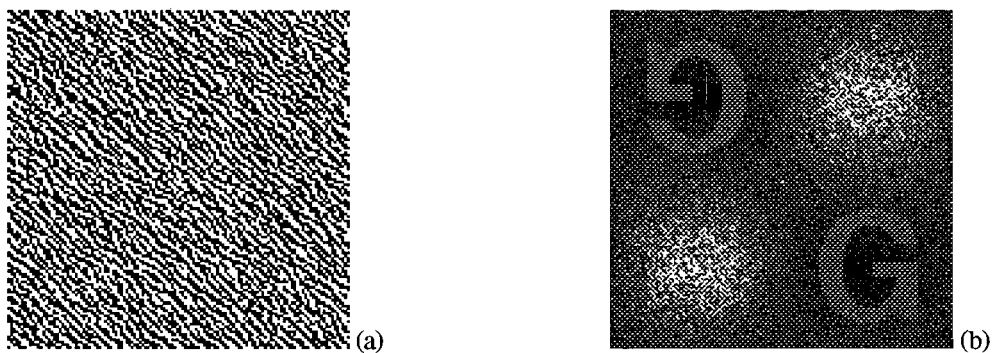


Fig. 6. Binary hologram generated using the error diffusion algorithm suggested by [9] (a), and its simulated reconstruction (b).

Because the algorithm suggested by [9] does not take into account of the size of the reconstruction region, it works well only for the reconstruction of small size objects.

In summary, the algorithm combining the ordinary error diffusion method with optimal diffusion coefficients (Eqs. (16) and (17)) can generate high quality binary holograms compared to other algorithms based on error diffusion. High quality means here, the error in the reconstruction region is small, the diffraction efficiency is relative high. Another advantage of this algorithm is its stability.

More tests have been done for the algorithm proposed here [18]. The result indicates that a better reconstruction can be achieved if an appropriate diffusion size is used, but the improvements are limited [18]. We conclude that the computational efficiency of error diffusion, which is $O(N)$ (N is the size of the hologram), has to be contrasted to its limited reconstruction quality for intensity signals.

6. Iterative method

As indicated in sections (2) and (4), the convergence of the iterative binarization algorithm depends on the enhancement factor $\beta^{(i)}$ if the number of binarized pixels is smaller than N^2 . In the following, two approaches are suggested to determine $\beta^{(i)}$. The strategy is that the signal in region w is enhanced proportionally to the strength of the binarization noise in this region.

The binarization process can be performed in several steps [14,16], say S steps. In step s ($s=1, 2, \dots, S$), K_s pixels are binarized. The binarization process is performed in this step until the stopping criterion is satisfied. Then the $(s+1)$ -th step is started. In this step K_{s+1} ($K_{s+1} > K_s$) pixels are binarized etc.. Each step results in a reduction of MSE except the last step, in which all pixels are binarized to generate the final binary hologram. If MSE does not change very much with iterations, the hologram distribution arrives at a steady state.

As an example, Fig. 4 is used again here for the iterative binarization algorithm. Other parameters remain unchanged. We use the iterative algorithm Fig. 1, which is performed in 2 steps ($S=2$). The algorithm is controlled by the number of iterations. To see the behavior of MSE more clearly, 24 iterations are performed in the first step, in which $K_1=12288$ pixels (which is $3/4$ of all hologram pixels N^2) are chosen randomly and binarized. In the second step, or the last step, $K_2=16384$ or all N^2 pixels are binarized to get the final binary hologram.

Approach one: $\beta_1^{(i)}$ is chosen according to the following criterion: the signal in region w is enlarged by factor $(1+\beta_1^{(i)})$, so that the energy in this region, $H_{(x,y)} |(1+\beta_1^{(i)})h^{(i)}(x,y)|^2$, with respect to that of the signal, $H_{(x,y)} |h^{(i)}(x,y)|^2 = H_{(x,y)} |h^{(i)}(x,y) + e^{(i)}(x,y)|^2$, remains conservative. Here $h^{(i)}(x,y)$ is the reconstruction of the binarized hologram. Thus we have

$$\beta_1^{(i)} = \sqrt{\frac{\sum_w |h^{(i)}(x,y) + e^{(i)}(x,y)|^2}{\sum_w |h^{(i)}(x,y)|^2}} - 1 \quad (19)$$

Mapping all terms in $\beta_1^{(i)}$ into hologram domain (referred to (13)), we have

$$\beta_1^{(i)} = \sqrt{\frac{H^{(i)}CH^{(i)} + E^{(i)}CE^{(i)} + H^{(i)}(C^* + C)E^{(i)}}{H^{(i)}CH^{(i)}}} - 1 \quad (20)$$

where ACB is defined as in (14). (19) is used in algorithm Fig. 1, (20) in algorithm Eq. (10).

Fig. 7(a) shows the binary hologram generated using the iterative binarization algorithm Fig. 1 and enhancement factor $\beta_1^{(i)}$. The simulated reconstruction of Fig. 7(a) is shown in Fig. 7(b), where $MSE=0.0462$, $MSE^*=0.0347$, and the diffraction efficiency $\eta=5.03\%$.

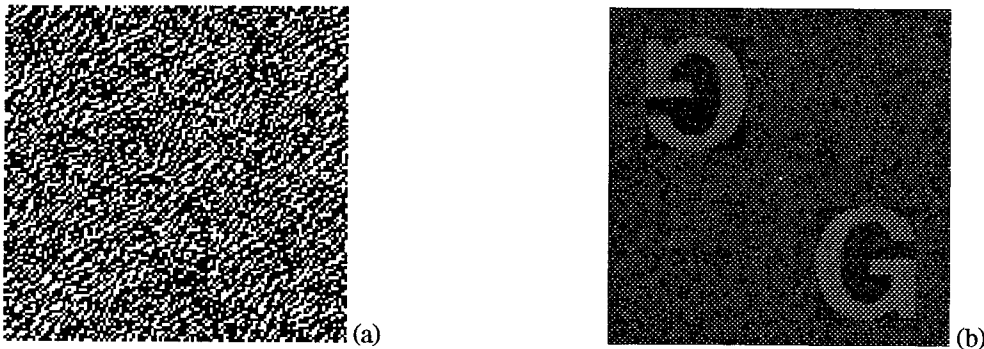


Fig. 7. Binary hologram generated using iterative binarization algorithm Fig. 1 and $\beta_1^{(i)}$ determined by approach one (a), and its simulated reconstruction (b).

Fig. 8 below shows the behavior of MSE with iterations for this example.

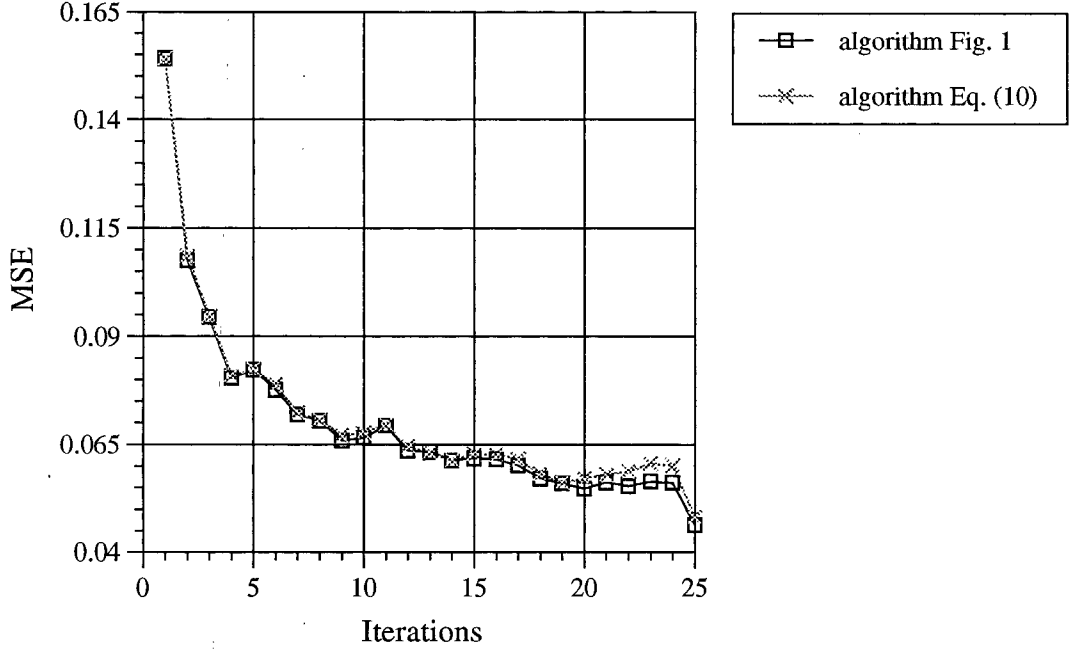


Fig. 8. MSE changes with iterations for $\beta_1^{(i)}$ of two algorithms, Fig. 1 and Eq. (10).

Two MSE curves are shown in Fig. 8. One (dark line with square symbol) is achieved by using the algorithm Fig. 1, namely changing from hologram domain to reconstruction domain and inversely, and the other (gray line with cross symbol) by using the algorithm Eq. (10), namely diffusing binarization error in hologram domain. It shows a good agreement of both. Note that $\beta_1^{(i)}$ given by (20) for algorithm Eq. (10) is computed here by a periodic correlation to save the computing time. Hence, overlaps in two sides of the correlation function appear, which give rise to the difference of two MSE curves above.

Obviously, the algorithm Eq. (10) only makes sense for analyses, not for the computation.

Approach two: $\beta_2^{(i)}$ is chosen according to the following criterion: the signal in region w is enlarged by factor $(1+\beta_2^{(i)})$, so that MSE of $(1+\beta_2^{(i)})h^{(i)}(x,y)$ with respect to signal $h^{(i)}(x,y) = h^{(i)}(x,y) + e^{(i)}(x,y)$ is minimum. Here $h^{(i)}(x,y)$ is the reconstruction of the binarized hologram.

Referred to the definition of factor α in MSE of (2), we have here

$$\beta_2^{(i)} = \frac{\sum_w |e^{(i)}(x,y)|^2 + \sum_w h^{(i)*}(x,y)e^{(i)}(x,y)}{\sum_w |h^{(i)}(x,y)|^2 + \sum_w h^{(i)}(x,y)e^{(i)*}(x,y)} \quad (21)$$

or in hologram domain

$$\beta_2^{(i)} = \left| \frac{E^{(i)}CE^{(i)} + H^{(i)}C^*E^{(i)}}{H^{(i)}CH^{(i)} + H^{(i)}CE^{(i)}} \right| \quad (22)$$

where ACB is defined as in (14).

The generated binary hologram of Fig. 4(b), using the iterative algorithm Fig. 1 and $\beta_2^{(i)}$, is shown in Fig. 9(a). Its simulated reconstruction is shown in Fig. 9(b), where $MSE=0.0450$, $MSE^{\#}=0.0332$, $\eta=5.46\%$.

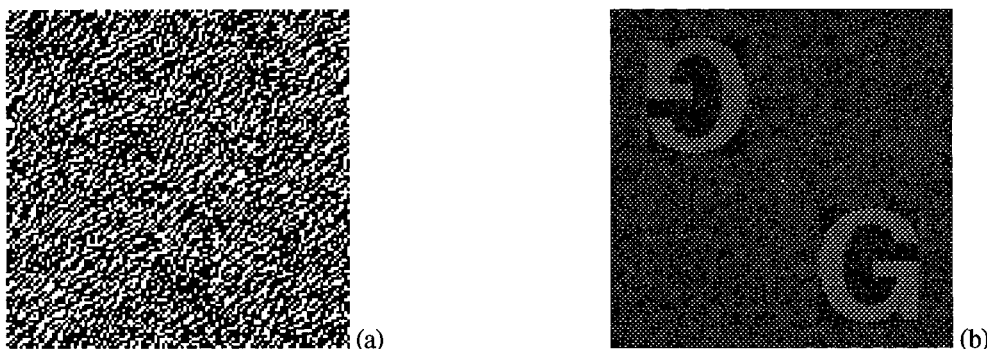


Fig. 9. Binary hologram generated using iterative binarization algorithm Fig. 1 and $\beta_2^{(i)}$ determined by approach two (a), and its simulated reconstruction (b).

The behaviors of MSE with iterations for the example above are shown in Fig. 10. Five MSE curves were presented for different $\beta^{(i)}$. Each curve is the average of MSE from five binarization processes with different initial conditions. The first one (dark line with square symbol) shows the changes of MSE for $\beta = \beta_1^{(i)}$, $\beta_1^{(i)}$ is determined by approach one; the second one (gray line with cross symbol) for $\beta = \beta_2^{(i)}$, $\beta_2^{(i)}$ is determined by approach two; the third one (gray line with circle symbol) for $\beta = 0.8\beta_2^{(i)}$, etc. The results show the dependence of the convergence rate and the error reduction degree of the algorithm on $\beta^{(i)}$.

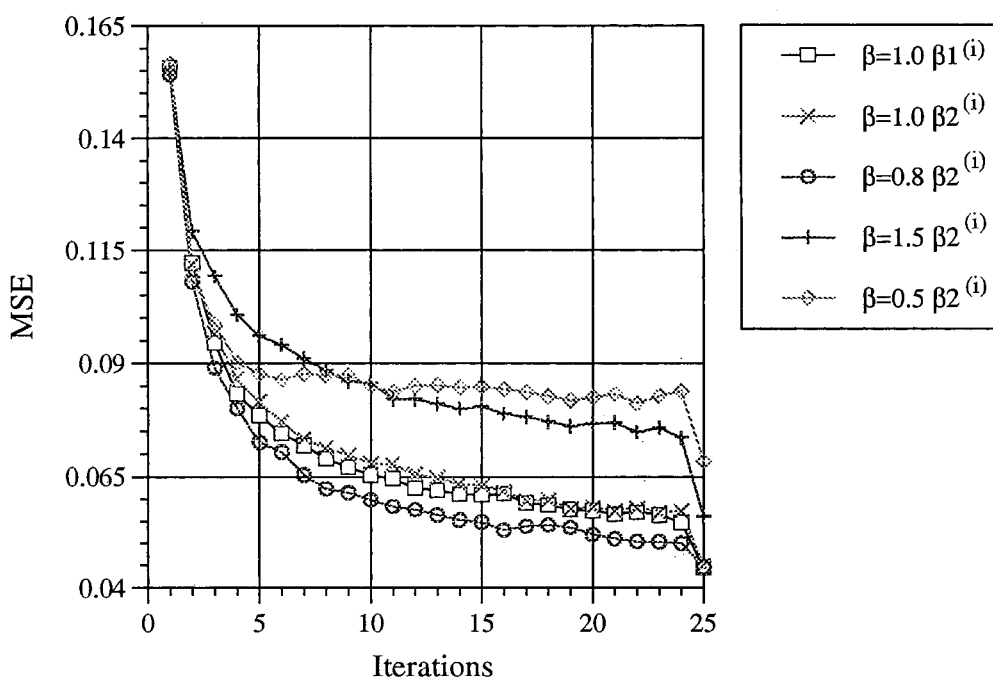


Fig. 10. MSEs change with iterations for different $\beta^{(i)}$.

There is little difference in convergence rate and error reduction degree for the two approaches determining $\beta^{(i)}$. However, if an appropriate $\beta^{(i)}$ is used, e.g. $\beta = 0.8\beta_2^{(i)}$ in Fig. 10, the convergence is really speeded up. In contrast, inappropriate $\beta^{(i)}$ will change the convergence rate, and therefore affect the error reduction degree (see $\beta = 1.5\beta_2^{(i)}$ and $\beta = 0.5\beta_2^{(i)}$ in Fig. 10).

In summary, iterative binarization algorithm can be used to applications where high reconstruction quality of complex signals, and high diffraction efficiency are required. It is also applicable for reconstruction of intensity signals, if one step in algorithm Fig. 1 is changed, i.e. instead of correcting complex values in the reconstruction region, only the amplitude is corrected. $MSE^{\#}$ defined in (18) is computed. For reconstruction of intensity signals, higher

reconstruction quality and higher diffraction efficiency will be achieved (the example of Fig. 4(b) above gives $MSE=0.0263$, $\eta=7.20\%$). The computational complexity of iterative binarization algorithm is $O(N\log_2 N)$, where N is the size of the hologram.

Conclusions

A rigorous mathematical analysis of computer generated binary Fourier transform holograms was performed. By detailing the binarization process step by step, a generalized error diffusion coefficients matrix was derived. Algorithms for generation of binary holograms, which are normally divided into noniterative methods and iterative methods, were discussed, their properties and efficiencies were investigated. In the view of the analysis results we conclude noniterative methods and iterative methods are the same in principle. They are distinguished only by how to diffuse the binarization error, at one time or at several times. The examples presented show that the combination of the ordinary error diffusion method with optimal diffusion coefficients is a valuable tool for binary Fourier transform holography. However, it is only applicable for reconstruction of intensity signals. Besides, the diffraction efficiency from such holograms is lower. In contrast, the iterative binarization algorithm with random binarization of hologram pixels has demonstrated the ability to find the minimum (local or global) of MSE. It can be used for applications where complex signals, or intensity signals are required.

References

- [1] Wilfrid B. Veldkamp and Thomas J. McHugh, "Binary Optics", Scientific American (1992) pp. 92-97.
- [2] Andrew J. Lee and David P. Casasent, "Computer generated hologram recording using a laser printer", Applied Optics, Vol. 26, No. 1 (1987) pp. 136-138.
- [3] T. C. Strand "Signal/noise in analog and binary holograms", Opt. Eng. 13 (1974) pp. 219-227.
- [4] R. W. Floyd and L. Steinberg, "An adaptive algorithm for spatial grayscale", Proc. Soc. Inf. Disp. Vol. 12 (1976) pp. 55-77.
- [5] Richard Hauck and Olof Bryngdahl, "Computer-generated holograms with pulse-density modulation", J. Opt. Soc. Am, A / Vol. 1, No. 1 (1984) pp. 5-10.
- [6] S. Weissbach, F. Wyrowski and O. Bryngdahl, "Quantization noise in pulse density modulated holograms", Optics communications, Vol. 67, No. 3 (1988) pp. 167-171.
- [7] Etienne Barnard, "Optimal error diffusion for computer-generated holograms", J. Opt. Soc. Am, A / Vol. 5, No. 11 (1988) pp. 1803-1871.
- [8] Reiner Eschbach, "Comparison of error diffusion methods for computer-generated holograms", Applied Optics / Vol. 30, No. 26 (1991) pp. 3702-3710.
- [9] Severin Weissbach and Frank Wyrowski, "Error diffusion procedure: theory and applications in optical signal processing", Applied Optics / Vol. 31, No. 14 (1992) pp. 2518-2534.
- [10] Reiner Eschbach and Zhigang Fan, "Complex-valued error diffusion for off-axis computer-generated holograms", Applied Optics / Vol. 32, No. 17 (1993) pp. 3130-3136.
- [11] Michael A. Seldowitz, Jan P. Allebach, and Donald W. Sweeney, "Synthesis of digital holograms by direct binary search", Applied Optics / Vol. 26, No. 14 (1987) pp. 2788-2798.
- [12] Brian K. Jennison and Jan p. Allebach and Donald w. Sweeney, "Efficient design of direct-binary-search computer-generated holograms", J. Opt. Soc. Am. A / Vol. 8, No. 4 (1991) pp. 652-660.
- [13] Myung Soo Kim and Clark C. Guest, "Block-quantized binary-phase holograms for optical inter-connection", Applied Optics / Vol. 32, No. 5 (1993) pp. 678-683.
- [14] Frank Wyrowski, "Iterative quantization of digital amplitude holograms", Applied Optics / Vol. 28, No. 18 (1989) pp. 3864-3869.
- [15] Chien-Hsion Wu, Chia-Lun Chen, and M. A. Fiddy, "Iterative procedure for improved computer-generated-hologram reconstruction", Applied Optics / Vol. 32, No. 26 (1993) pp. 5135-5140.
- [16] Eryi Zhang, Steffen Nohte, Christoph H. Dietrich, and Reinhard Männer, "Gradual and random binarization of gray scale holograms", accepted by Applied Optics (1995).
- [17] R.W. Gerchberg and W.O. Saxton, "A practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures", Optik / Vol. 35, No. 2 (1972) pp. 237-246.
- [18] Eryi Zhang, Steffen Nohte, Christoph H. Dietrich, Reinhard Männer, "Half-optimal Error Diffusion for Binary Fourier Transform Holograms", submitted to J. Opt. Soc. Am.A (1995).